Rc^2: R collaboration in the cloud

E. James Harner[1⋆] **and Mark Lilback**[1]

1. Department of Statistics, West Virginia University
⋆Contact author: jharner@stat.wvu.edu

Keywords: R, cloud computing, collaboration, web 2.0

Rc^2 is a cloud-based, collaborative, web 2.0 interface to R that works with WebKit-based browsers (e.g., Safari and Chrome). In addition to desktop browsers, client-specific style sheets and scripting provide a touch-optimized interface for mobile devices such as the iPad. With Rc^2, R sessions are no longer tied to a specific computer or user.

Rc^2 is designed for simplicity and ease-of-use from the start, allowing students to learn R without solely imposing a command-line interface. At the same time, power users will find the system flexible enough to meet most of their needs. Users can have per-project workspaces, and instructors can predefine workspaces containing the required data and R packages for each assignment.

The groundbreaking features of Rc^2 arise from the collaborative and instructional benefits that cloud-based computing brings. Instructors can schedule classroom sessions where students can watch the instructor interact with R in real-time. The instructor can turn control of R over to individual students as a virtual blackboard to pose questions and work through problems while still communicating via voice chat. Integration with the social web allows instructors to provide notifications to students over the messaging platform of their choice, be it email, SMS, Twitter, or Facebook.

The same features allows researchers to collaborate over the Internet without concern for data becoming out of sync. Users can start long-running computations and Rc^2 will notify the user(s) when the process is complete. Full support for Sweave allows users to easily include, update, and format R output within LaTeX documents for both classroom assignments and publishable papers.

Rc^2 is designed to be a fast, scalable, distributed system with features like load-balancing, multiple authentication mechanisms (password, LDAP, Kerberos), dynamic auto-configuration using mDNS and DNS-SD, fast and persistent client-server communications using WebSockets, and custom R packages for database interaction and graphics generation. Rc^2 can start on a single server and expand to a cluster of computers optimized for each tier of the system (web server, application server, R instances, and database).